

# **Bases de données avancées**

## **—**

# **Introduction BD relationnelles**

Bruno PINAUD

[bruno.pinaud@u-bordeaux.fr](mailto:bruno.pinaud@u-bordeaux.fr)

**SUPPORT DE COURS VOLONTAIREMENT INCOMPLET**

### **Bibliographie**

Georges Gardarin, Base de données, Ed. Eyrolles, 2003

A First Course in Database Systems (3rd edition) J. D. Ullman, J. Wildom, Ed.

Pearson, 2008



# INTRODUCTION ET DÉFINITIONS



# *Définition : base de données*

- Plus qu'un simple ensemble de données non-indépendantes
- Permet l'indépendance données – programme

« [...] Ensemble de données modélisant les objets d'une partie du monde réel et servant de support à une application informatique » G. Gardarin

## **Propriété importante**

interrogeable par le contenu selon n'importe quel critère

Exemple : quels sont les numéros de Radioactive Man à moins de 10\$ ?

# *Introduction*

*Processus de construction d'une base de données*

# Définition : SGBD Relationnel

- **SGBD** : Système de Gestion de Bases de Données (parfois SGBDR)
- **DBMS** : DataBase Management System (parfois RDBMS)
- Outil informatique pour gérer des BD.
- Régie par le **Modèle relationnel** (Codd, 70)
- **Doit absolument posséder plusieurs fonctionnalités**
  - ✓ Sauvegarde (persistance) des données
  - ✓ Interrogation (SQL) des données
  - ✓ Recherche et mise en forme des données stockées
  - ✓ Partage des données entre les différents utilisateurs
  - ✓ Gestion de la concurrence d'accès
  - ✓ Sécurité des données (gestion des incidents)
  - ✓ Optimisation des opérations dans un souci constant de performance

# Définition : SGBD Relationnel

Les principales opérations (LDD, LMD et LCD, SQL-92) :

- Création d'une relation (*create table*)
- Suppression d'une relation (*drop table*)
- Modification de la définition d'une relation (*alter table*)
  
- Interrogation (*select*)
- Insertion (*insert*)
- Mise à jour (*update*)
- Suppression (*delete*)
  
- Accorder des droits (*grant*)
- Retirer des droits (*revoke*)
- Validation d'une transaction (*commit*)
- Annulation d'une transaction en cours (*rollback*)

# Définition : SGBD Relationnel

Les principales opérations (LMD) :

- Interrogation (*select*)
- Insertion (*insert*)
- Mise à jour (*update*)
- Suppression (*delete*)

sont regroupées dans des **transactions**.

Une **transaction** est un regroupement atomique d'un ensemble d'opérations.

Un **Système de Gestion de Bases de Données** (SGBD) se charge de la mise en œuvre des transactions.

Régies par le **Modèle relationnel** (Codd, 70) : tous les SGBD fonctionnent à peu près de la même façon (au moins pour le LDD et LMD) : « One Size Fits All ».

**Propriétés ACID : Atomicité, Cohérence, Isolation, Durabilité**

# ***BD : Vocabulaires et définitions***



# *Concepts de base : domaine*

## **Définition**

- Ensemble de valeurs atomiques (non décomposable)
- Équivalent au typage en programmation

# Concepts de base : produit cartésien

## Définition

Le produit cartésien d'un ensemble de domaines  $D_1, D_2, \dots, D_n$  noté  $D_1 \times D_2 \times \dots \times D_n$  est l'ensemble des n-uplets (tuples)  $\langle v_1, v_2, \dots, v_n \rangle$  tels que  $v_i \in D_i$ .

# *Concepts de base : relation*

## **Définition**

- Sous ensemble  $r$  du produit cartésien d'un ensemble de domaines
- Caractérisée par un nom

# *Concepts de base : relation*

## **Définition**

- Sous ensemble  $r$  du produit cartésien d'un ensemble de domaines
- Caractérisée par un nom

# *Concepts de base : relation*

## **Définition**

- Sous ensemble  $r$  du produit cartésien d'un ensemble de domaines
- Caractérisée par un nom

# *Concepts de base : attribut*

## **Définition**

- Nom donné à une colonne
- Composé d'un identifiant et d'un domaine
- Nombre d'attributs d'une relation = degré de la relation (ou arité)

# Concepts de base : schéma de relation

## Définition

- Un **schéma de relation**  $R$  est défini par un ensemble d'attributs  $U$  et un ensemble de contraintes.
- On le note couramment  $R(U)$ .
- Le schéma décrit l'**intention** de la relation.
- La relation (tableau) définit une **extension**.
- Une relation  $r$  est une instance finie d'un schéma de relation, notée  $r:R(U)$ .

# Concepts de base : base de données

## Définition

- Un **schéma de base de données** est un ensemble de schémas de relations **liés** par des **dépendances référentielles** : attributs communs ou plus généralement des dépendances d'inclusion.
- Une **base de données** est alors un ensemble de relations (extensions) associé au schéma de base de données et vérifiant ses **contraintes d'intégrité**.



# **INTRODUCTION AU MODÈLE RELATIONNEL**



# *Un mauvais exemple*

**Objectif** : Modélisation d'une compagnie aérienne.

On veut savoir quel est l'avion utilisé pour chaque vol et sa capacité.

Point de départ : création d'une **relation universelle** avec tous les attributs

## **Transports**

Vol	Avion	Capacité
IT5033	Airbus A330	335
AF2401	Airbus A330	335
AF2409	Boeing 727	150
IT5133	Airbus A330	335
IT5035	Canadair	90
AF2802	Airbus A330	335

# *Un mauvais exemple*

## **Transports**


Vol	Avion	Capacité
IT5033	Airbus A330	335
AF2401	Airbus A330	335
AF2409	Boeing 727	150
IT5133	Airbus A330	335
IT5035	Canadair	90
AF2802	Airbus A330	335

Anomalies de la relation  
**Transports**

# *Le modèle relationnel*

- Défini par E. F. Codd (IBM Research) dans « A Relational Model of Data for Large Shared Data Banks », CACM 13, No. 6, June 1970)
- Indépendant de la représentation physique des données
- Assise mathématique forte (algèbre relationnelle, formes normales)

# Conception d'un schéma relationnel

**But** : Éliminer les anomalies de la relation universelle pour faciliter la manipulation des relations  **Normaliser les relations.**

**Méthode** : **Décomposer** la relation universelle en sous-relations qui ne souffrent pas des anomalies précédentes.

**Problématique** : obtenir un nouveau schéma de base de données qui :

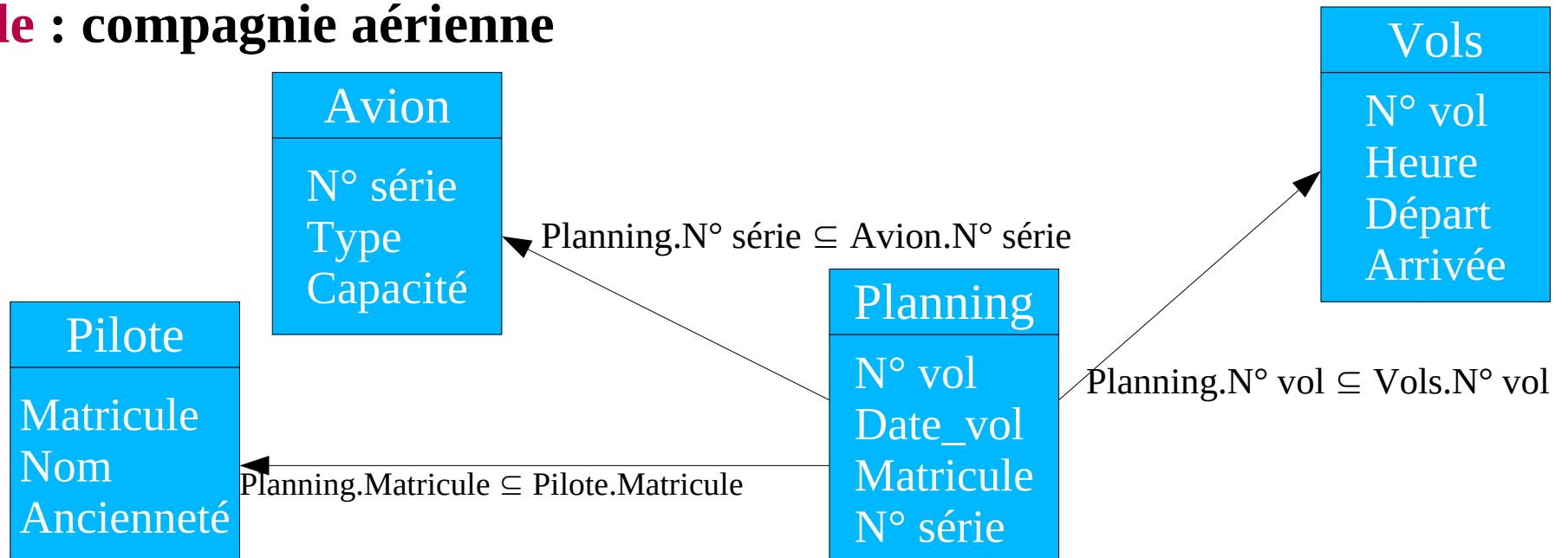
- ✓ conserve toutes les données,
- ✓ conserve un minimum de contraintes d'intégrité mais respecte le réel perçu,
- ✓ élimine toutes les anomalies.

# Concepts de base : base de données

## Définition (rappel)

- Un **schéma de base de données** est un ensemble de schémas de relations **liés** par des **dépendances référentielles (un type de contraintes)** : attributs communs ou plus généralement des dépendances d'inclusion.
- Une **base de données** est alors un ensemble de relations (extensions) associé au schéma de base de données et vérifiant toutes ses **contraintes**.

## Exemple : compagnie aérienne



# Concepts de base : base de données

**Avion**

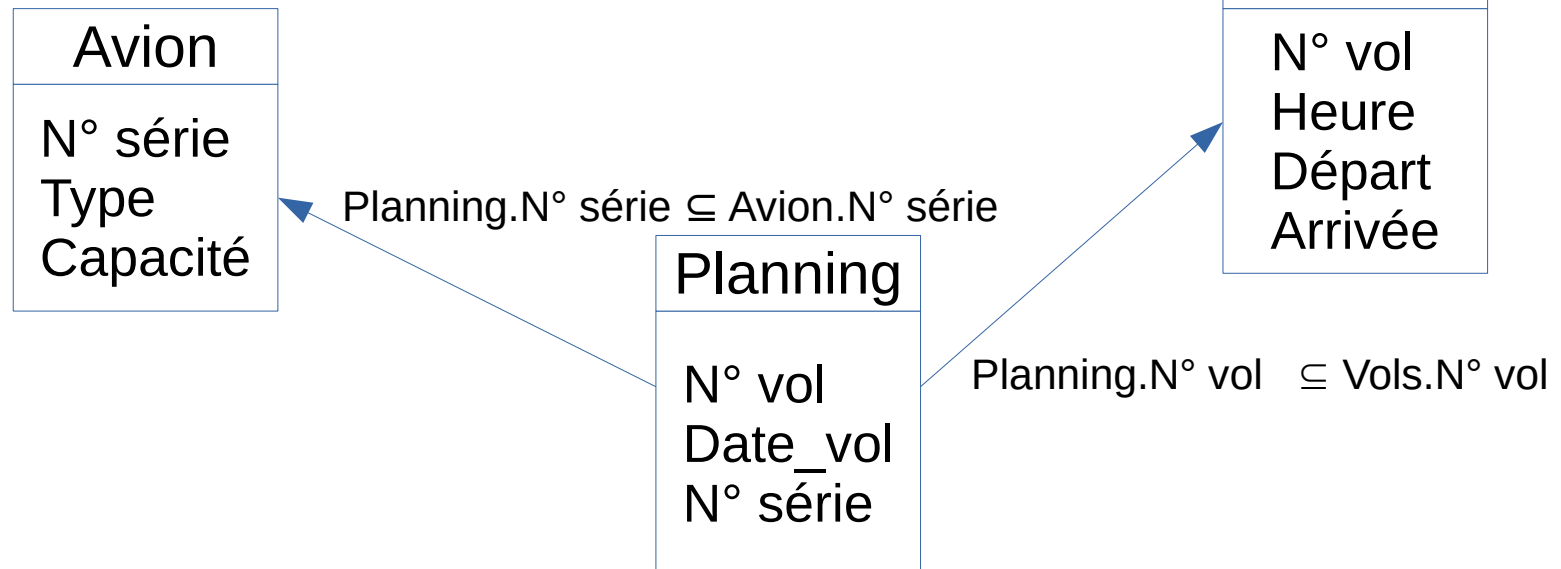
No Série	Type	Capacité
2001	Airbus A330	335
2002	Airbus A330	335
301	Boeing 727	150
94	Canadair	90

**Planning**

No Vol	Date_vol	No série
IT5033	...	2001
AF2401		2002
AF2409		2002
IT5133		301
IT5035		301
AF2802		2002

**Vol**

No Vol	Heure Départ	Heure Arrivée
IT5033	...	
AF2401		
AF2409		
IT5133		
IT5035		
AF2802		



# Contraintes d'intégrité

## Définition

- Tout schéma de base de données doit être conçu pour imposer le respect d'un maximum de contraintes d'intégrité du réel perçu.
- Les **contraintes d'intégrité** sont des expressions logiques qui **doivent être satisfaites à tout instant** par une instance de base de données.
- Plusieurs types
  - ✓ Contraintes sur les attributs : domaines, valeurs nulles, ... ;
  - ✓ Contraintes sur les  $n$ -uplets : la valeur d'un attribut peut dépendre d'un ou plusieurs autres attributs du même  $n$ -uplet (dates d'emprunt et de retour pour une bibliothèque), ... ;
  - ✓ Contraintes sur les relations : clés, cardinalité, ... ;
  - ✓ Contraintes sur la base de données : clés étrangères, ... ;
  - ✓ Contraintes temporelles : évolution chronologique (diplômes, état-civil), ...

**Problématique : minimiser le nombre de contraintes tout en restant équivalent à l'ensemble des contraintes d'origine**



# Contraintes d'intégrité : notion de clé

- Une relation est un ensemble de  $n$ -uplets. Par définition, un ensemble n'a pas d'élément en double, donc chaque  $n$ -uplet d'une relation est unique.
- Pour identifier les  $n$ -uplets de façon unique sans en donner toutes les valeurs et respecter leur unicité une clé est nécessaire.
- **Clé** : groupe minimum d'attributs qui détermine chaque  $n$ -uplet de façon unique.

# Contraintes d'intégrité : notion de clé

## Propriétés

- Toute relation possède au moins 1 clé : l'ensemble de ses attributs
- Si une relation possède plusieurs **clés candidates**, on choisit une clé qui sera privilégiée : **la clé primaire**. Aucun des attributs d'une clé primaire n'admet de valeur nulle (vide).

## Conventions d'écriture

- On souligne la clé primaire.
- Les clés candidates sont soulignées en pointillés (souvent omis).

# Contraintes d'intégrité : notion de clé

## Exemple

Soit la relation  $R(A,B,C,D,E)$  :

A	B	C	D	E
A1	B1	C1	D1	E1
A1	B1	C2	D1	E1
A1	B1	C5	D1	E1
A1	B2	C5	D4	E1
A2	B1	C2	D1	E1
A2	B1	C5	D1	E1
A2	B1	C5	D3	E1
A2	B2	C5	D4	E1
A3	B3	C5	D1	E2

Est-ce que (A, C, D) est une clé candidate de R ?